

## ロタウイルス感染症の血清疫学的データの対数線型モデルによる解析

秋 葉 澄 伯

札幌医科大学衛生学講座 (主任 浦沢正三 教授)

### An Analysis of Seroepidemiological Data of Rotavirus Infection by Loglinear Models

Suminori AKIBA

Department of Hygien, Sapporo Medical College

(Chief : Prof. S. Urasawa)

We have developed a loglinear analysis program for personal computers (microcomputers) and have used it in the analysis of age distributions of neutralizing antibody to rotavirus of inhabitants in Sapporo and in six cities in Southeast Asia. The following results were obtained:

- 1) A loglinear model with no three factor effect was found to fit the antibody distributions in the six Southeast Asian cities by an analysis using three factors (antibody titer, age and city).
- 2) Utilizing this model, the age distributions of antibody in these cities were grouped into three.
- 3) The antibody distribution in Sapporo was so different from those of the Southeast Asian cities that a loglinear analysis could not be carried out.

- 4) This loglinear analysis program for the personal computers is easy to handle and should be useful in epidemiological studies.

(Received December 1, 1982 and accepted January 31, 1983)

**Key words:** Rotavirus infection, Seroepidemiology, Loglinear analysis, Microcomputer

## 1 結 言

3元以上の分割表の解析には対数線型解析が有用であり、疫学の分野でもこれを用いた研究が増えている。著者は最近パーソナルコンピュータによる対数線型プログラムを開発し、これを用いて東南アジアおよび日本のロタウイルスに対する年齢別抗体価の検討を行ったので報告する。

## 2 方 法

インドネシア5都市 (Samarinda, Kolaka, Surabaya, Lombok, Jayapura), ラオス(Vientien), 日本(札幌)計7都市住民の抗ロタウイルス中和抗体価を牛ロタウイルス NCDV (Neonatal Calf Diarrhea Virus), LLCMK<sub>2</sub>細胞を用いたマイクロプレート法により測定した。方法の詳細については Urasawa *et al.*<sup>1)</sup>を参照されたい。

対数線型解析プログラムは NEC のパーソナルコンピュータ PC-8001 を用い、N-BASIC で作成した。ここではプログラムそのものについての説明は省略し、

対数線型解析の基本的原理と、著者の年齢別抗体価測定成績へのその適用について述べる。(以下の説明は、Fienberg<sup>2)</sup>, Upton<sup>3)</sup>, Anderson *et al.*<sup>4)</sup>らの著書を参考にしてまとめたものである。)

Table 1 は、札幌市に於ける年齢別の抗ロタウイルス中和抗体の有無(抗体価 $\leq 4\times$ を抗体陰性とした。)を示したものである。年齢と中和抗体陽性率が独立であるという仮定のもとで期待値を求めるには次式による計算を行う。

$$E_{ij} = \frac{X_{i.} \times X_{.j}}{X_{..}} \dots\dots\dots (1)$$

$i=1, 2$

$j=1, 2$

$X_{ij}$ : 観測値

$E_{ij}$ : 期待値

$X_{..}$ : 総計

$X_{i.}, X_{.j}$ : 周辺計

} Table 1 を参照

得られた期待値を用いて通常の如く<sup>5)</sup>,  $X^2$  を計算すると

$$X^2 = 93.9 \quad \text{自由度} = 1$$

**Table 1** Prevalence of neutralizing antibody to rotavirus of inhabitants in Sapporo

Age	Antibody		Total
	— (4×)	+ (≥8×)	
0-9	145 (X <sub>11</sub> )	111 (X <sub>12</sub> )	256 (X <sub>1.</sub> )
10-	54 (X <sub>21</sub> )	258 (X <sub>22</sub> )	312 (X <sub>2.</sub> )
Total	199 (X <sub>.1</sub> )	369 (X <sub>.2</sub> )	568 (X <sub>..</sub> )

This table was made from Table 3 to be used in the explanation of the loglinear analysis under "Method."

$$p < 0.01$$

となり年齢と中和抗体陽性率が独立であるとする仮定は否定される。

(1)式の両辺について自然対数をとると

$$\begin{aligned} \log E_{ij} &= \log X_{i.} + \log X_{.j} - \log X_{..} \\ &= \log X_{..} + (\log X_{i.} - \log X_{..}) \\ &\quad + (\log X_{.j} - \log X_{..}) \\ &= GM + T_i + A_j \end{aligned}$$

但し,

$$\begin{aligned} GM &= \log X_{..} \\ T_i &= \log X_{i.} - \log X_{..} \\ A_j &= \log X_{.j} - \log X_{..} \end{aligned}$$

と書くことができる。log E<sub>ij</sub> は周辺計の総平均からの偏りを表わすパラメーター T<sub>i</sub>, A<sub>j</sub> と総平均を表わすパラメーター GM の和で表わされている。この式は一つの対数線型モデルを表わしているが、前述の X<sup>2</sup>検定の結果、年齢と中和抗体陽性率が独立であるとする仮定は否定されており、そのような仮定のもとに作られたこのモデルもまた、観察されたデータをよく説明していないことは明らかである。適合度をよくするため新たなパラメーターを加える。

$$\log E_{ij} = GM + T_i + A_j + (T \cdot A)_{ij} \dots\dots\dots(2)$$

新しく加わった(T·A)<sub>ij</sub> は2因子交互作用を表わしており、抗体陽性率が年齢によってどのような変化を受けるかを表わす。このモデルは飽和モデルであり、データへのモデルの適合は完全である。即ち、このモデルで得られる対数期待値は対数観測値と一致する。なお、対数線型モデルのパラメーターには下式であらわされるような総和をゼロとする制約がある。

$$\begin{aligned} \sum_i T_i &= 0 & \sum_j A_j &= 0 \\ \sum_i (T \cdot A)_{ij} &= 0 & \sum_j (T \cdot A)_{ij} &= 0 \end{aligned}$$

従って

$$\begin{aligned} T_1 &= T_2 \\ A_1 &= A_2 \end{aligned}$$

$$(T \cdot A)_{11} = (T \cdot A)_{21} = (T \cdot A)_{12} = (T \cdot A)_{22}$$

となる。この結果、独立なパラメーターの数は、

$$\begin{aligned} &1 \quad (\rightarrow GM \text{ 由来}) \\ &+2 \quad (\rightarrow T_1, T_2 \text{ 由来}) \\ &+2 \quad (\rightarrow A_1, A_2 \text{ 由来}) \\ &+4 \quad (\rightarrow (T \cdot A)_{11}, (T \cdot A)_{21}, (T \cdot A)_{12}, \\ &\quad (T \cdot A)_{22} \text{ 由来}) \end{aligned}$$

$$= 9 \text{ 個}$$

ではなく

$$1+1+1+1=4 \text{ 個}$$

である。この数は2×2表のセルの数4個と一致するが、このような関係は他の分割表の対数線型モデルについても成り立つ。(但し、セルにゼロが出現する場合、調整の必要なことがある。詳しくは Fienberg<sup>2)</sup>を参照されたい。)

Table 2 のような3元分割表に対応する飽和モデルは次式で表わされる。

$$\begin{aligned} \log E_{ijk} &= GM + T_i + A_j + C_k \\ &\quad + (T \cdot A)_{ij} + (A \cdot C)_{jk} + (T \cdot C)_{ik} \\ &\quad + (T \cdot A \cdot C)_{ijk} \\ i &= 1, 2 \\ j &= 1, 2 \\ k &= 1, 2 \\ T &: \text{Antibody } (-, +) \\ A &: \text{Age} \\ C &: \text{City} \end{aligned}$$

**Table 2** Prevalence of neutralizing antibody to rotavirus of inhabitants in Sapporo and Jayapura

	Sapporo		Jayapura	
	Antibody			
Age	—	+	—	+
0-9	145	111	2	17
10-	54	258	5	96

This table was made from Table 3 to be used in the explanation of the loglinear analysis under "Method."

T<sub>i</sub>, A<sub>j</sub>, C<sub>k</sub> は主作用(因子)を表わすパラメーターで、各々抗体陽性者の数、年齢分布、両都市のサンプルサイズを反映している。(T·A)<sub>ij</sub>, (A·C)<sub>jk</sub>, (T·C)<sub>ik</sub> はいずれも2因子交互作用をあらわすパラメーターであって、各々、抗体陽性率の年齢による変化、両都市の年齢分布の違い、抗体陽性率の都市による違いを表わす。(T·A·C)<sub>ijk</sub> は3因子交互作用を表わすパラメーターであって、ここでは抗体陽性率の年齢による

変化の仕方が両都市でどのように異なるかを表わす。(パラメーターの意味については期待値の求め方の項目でも述べる。)

飽和モデルから任意のパラメーターを除けば不飽和モデルが得られる。しかし、ここでは階層モデルと呼ばれるもののみを考えるから、仮に2因子交互作用( $T, A$ )<sub>ij</sub>がモデルに含まれたなら、GM,  $T_i$ ,  $A_j$ など下位のパラメーターは全てモデルに含まなければならない。

与えられたモデルに基づいて期待値を求める方法を述べる。例として table 2 の分割表に次のモデルをあてはめる場合を考える。

$$\log E_{ijk} = GM + T_i + A_j + C_k + (T \cdot A)_{ij}$$

$$i=1,2$$

$$j=1,2$$

$$k=1,2$$

モデルに含まれるパラメーターは次のように対応する周辺計、総計を固定する働きをしている。

$$E_{..} = X_{..} \quad \dots\dots\dots (3)$$

$$E_{i.} = X_{i.} \quad \dots\dots\dots (4)$$

$$E_{.j} = X_{.j} \quad \dots\dots\dots (5)$$

$$E_{..k} = X_{..k} \quad \dots\dots\dots (6)$$

$$E_{ij.} = X_{ij.} \quad \dots\dots\dots (7)$$

上式の内、(6), (7)式が成立すれば(3)–(5)式は成立する。従って期待値は(6), (7)の2式を満たさねばならない。まず、初期値  $E^{(0)}_{ijk}$  全てに1を代入する。次に次式による計算を行って  $E^{(1)}_{ijk}$  を得て、

$$E^{(1)}_{ijk} = \frac{E^{(0)}_{ijk} \times X_{..k}}{E^{(0)}_{..k}}$$

更に次式による計算から、 $E^{(2)}_{ijk}$  を得る。

$$E^{(2)}_{ijk} = \frac{E^{(1)}_{ijk} \times X_{ij.}}{E^{(1)}_{ij.}}$$

この  $E^{(2)}_{ijk}$  は(6), (7)式を満足しており、このようにして求められた  $E^{(2)}_{ijk}$  が前述の対数線型モデルの期待値(最尤推定値)であることが知られている<sup>2)</sup>。3元表の対数線型モデルに基づく期待値はこのようにして簡単に求められるが、以下に示すモデルだけは例外である。

$$\log E_{ijk} = GM + T_i + A_j + C_k + (T \cdot A)_{ij} \\ + (A \cdot C)_{jk} + (T \cdot C)_{ik}$$

この場合、以下の3式が満たされるように期待値を求める。

$$E_{ij.} = X_{ij.} \quad \dots\dots\dots (8)$$

$$E_{.jk} = X_{.jk} \quad \dots\dots\dots (9)$$

$$E_{i..} = X_{i..} \quad \dots\dots\dots (10)$$

従って、

$$E^{(1)}_{ijk} = \frac{E^{(0)}_{ijk} \times X_{ij.}}{E^{(0)}_{ij.}} \quad \dots\dots\dots (11)$$

$$E^{(2)}_{ijk} = \frac{E^{(1)}_{ijk} \times X_{.jk}}{E^{(1)}_{.jk}} \quad \dots\dots\dots (12)$$

$$E^{(3)}_{ijk} = \frac{E^{(2)}_{ijk} \times X_{i..}}{E^{(2)}_{i..}} \quad \dots\dots\dots (13)$$

を次々に計算していけば良いが  $E^{(3)}_{ijk}$  は(8), (9)をもはや満たさない。従って(11)–(13)の計算を更にくり返し行い(2回目の計算では  $E^{(0)}_{ijk}$  の代わりに  $E^{(3)}_{ijk}$  を用いる。他も同様)、(8)–(10)を同時に充分満足する値が得られたところで、くり返し計算を終了する。

次にモデルの適合度の検定方法を述べる。これには上述の方法で求められた期待値と観測値から、通常の  $X^2$  検定を行う方法もあるが、一般には次のように尤度比統計量(Likelihood Ratio Statistics)を計算し、この分布が  $X^2$  分布で近似されることを利用して  $X^2$  検定を行う。

$$L.R.S. = 2 \sum_i \sum_j \sum_k \log \frac{X_{ijk}}{E_{ijk}}$$

L.R.S. : Likelihood Ratio Statistics

このときの自由度は以下のように求められる。

自由度 = 分割表のセルの数 – モデルに用いた独立なパラメーターの数

この値は、飽和モデルからパラメーターを除いていてそのモデルを作ったと仮定したとき、除かれた独立なパラメーターの数に一致する。従って、飽和モデルから3因子交互作用を除いたモデルを考えるとその自由度は、

$$(2-1) \times (2-1) \times (2-1) = 1$$

である。これが

$$2 \times 2 \times 2 = 8$$

とならないのはパラメーターの総和をゼロとする制約があるからである。なお、分割表のセルにゼロが出現する場合、期待値、自由度の計算の際、調整の必要な場合がある。(この点については Fienberg<sup>2)</sup>を参照)

次に各因子、交互作用の L.R.S., 自由度の求め方を述べるが、その前にモデル表記法を簡略化しよう。

今後、対数線型モデルの表記の際、抗体価、年齢、都市の3因子を各々、 $T, A, C$ で表わし、 $T \cdot A, A \cdot C, T \cdot C$ は各々の因子の間の2因子交互作用、 $T \cdot A \cdot C$ は3因子交互作用を表わすものとする。以上の表記法は前述したパラメーターの表記法と似ているが、パラメーターを表わすとき即ち、その値そのものを問題にするときは  $i, j, k$  などの添字付で表わし、因子(主作用)、交互作用などを意味するときは添字なしで表わす。

また、 $T \times A \times C$ ,  $T \times A + A \times C$ などは以下のモデルを表わす。

$$T \times A \times C = GM + T + A + C + T \cdot A + A \cdot C + T \cdot C + T \cdot A \cdot C$$

$$T \times A + A \times C = GM + T + A + C + T \cdot A + A \cdot C$$

但し、GM は総平均を表わす。これらの表記法は、GLIM (Generalized Linear Interactive Modelling)<sup>9)</sup> に於ける表記法に従ったものである。

3 因子交互作用  $T \cdot A \cdot C$  の L.R.S., 自由度を求めるには、飽和モデル (L.R.S.=0, 自由度=0) から、3 因子交互作用を除いたモデルを作り、その L.R.S. と自由度を計算して 2 つのモデル (飽和モデルと 3 因子交互作用のないモデル) の L.R.S., 自由度、各々の差をとると、これが 3 因子交互作用の L.R.S., 自由度である。2 因子交互作用  $T \cdot A$  の L.R.S., 自由度も、同様に 3 因子交互作用のないモデルから  $T \cdot A$  を除いて L.R.S., 自由度を求め、2 つのモデルの L.R.S., 自由度の差をとって得られる。

以上、対数線型解析について説明したが、本プログラムでは期待値、L.R.S., 自由度の計算を行う他、パラメーターの値も求めることができる。なお、このプログラムで得られる値が正しいことを確認するため Goodman の ECTA プログラム<sup>3)</sup> を用いて同一のデータを解析したところ同一の結果を得た。

### 3 結 果

Table 3 に東南アジア 6 都市および札幌市の年齢別抗ロタウイルス中和抗体価測定成績を示した。まず、東南アジア 6 都市について対数線型解析を行い、Table 4 に示す結果を得た。この表の見方を説明すると、Result 欄 1 例目 Effect には交互作用が示されており、その L.R.S., 自由度が 2 列目、3 列目に記載されている。4 列目 Model は交互作用の L.R.S., 自由度を計算する際に使ったモデルで 5 列目、6 列目にその L.R.S., 自由度が記載されている。1 列目の交互作用の L.R.S., 自由度はこのモデルと、これに交互作用を加えたモデルとの L.R.S., 自由度の差として求められた。Table 7, および 8 も表示は全く同じである。

Table 4 に示した通り、3 因子交互作用  $T \cdot A \cdot C$ 、2 因子交互作用  $A \cdot C$  が有意でない。しかし、 $A \cdot C$  について考えてみると各都市の年齢別のサンプルサイズは実験計画により予め定められてあったから、これは常にモデルに含めるべきである。従って観察された 3 元表に当てはめるべきモデルは 3 因子交互作用  $T \cdot A \cdot C$  の

Table 3 Age distribution of neutralizing antibodies to NCDV in seven localities

City	Age	Antibody					
		4×	8×	16×	32×	64×	128×
Vientien	0-9	0	4	5	14	6	5
	10-19	1	1	5	11	8	4
	20-	0	2	23	38	29	9
Samarinda	0-9	0	1	10	4	3	2
	10-19	0	8	7	7	8	1
	20-	0	6	18	32	21	10
Kolaka	0-9	1	4	2	5	7	3
	10-19	0	3	2	10	7	3
	20-	0	4	16	34	26	17
Surabaya	0-9	0	7	5	3	3	2
	10-19	1	2	8	15	2	2
	20-	1	3	36	34	32	15
Lombok	0-9	0	5	5	2	3	5
	10-19	0	4	4	9	4	4
	20-	1	2	23	28	15	5
Jayapura	0-9	2	3	9	2	1	2
	10-19	3	4	15	9	3	1
	20-	2	3	18	22	10	11
Sapporo	0-9	145	55	32	18	5	1
	10-19	16	19	15	9	7	2
	20-	38	82	72	33	17	2

みを欠いたモデル  $T \times A + A \times C + T \times C$  であり、このモデルに基づく期待値を Table 5 に示した。このモデルには 2 因子交互作用  $T \cdot C$  が含まれているから、6 都市の年齢別抗体分布は異なるものの、3 因子交互作用  $T \cdot A \cdot C$  が含まれていないから年齢の上昇に伴う抗体価の変化には都市間に Table 6 に示すような規則性が存在する。即ち、例として年齢 0-9 歳、10-19 歳の 2 階級に属し、抗体価が 8×, 32× のものを取り出して、2×2 分割表を作り、その OR (Odds Ratio) を計算するとその値は 6 都市で共通である。(まるめの誤差を除く)。この関係は他の年齢、抗体価の組み合わせをとっても成立する<sup>9)</sup>。

Fig. 1 に示した Table 5 の期待値に基づく 6 都市の抗体分布を比べてみると Samarinda, Surabaya, Lombok の抗体分布は良く似ており、Vientien, Kolaka も似かよっている。また Jayapura もどちらかというと前者の群に似ている。これらの点を確認するため Table 7, 8 の対数線型解析を行った。Table 7 は Vientien, Kolaka 2 都市の解析結果であるが 2 因子交

Table 4 Results of loglinear analysis

Factor	
T : Titer	$\leq 4 \times / 8 \times / 16 \times / 32 \times / 64 \times / 128 \times$
A : Age	0-9 / 10-19 / 20-
C ; City	Vientien / Samarinda / Kolaka / Surabaya / Lombok / Jayapura

Result					
Effect	L. R. S.	D. F.	Model	L. R. S.	D. F.
T . A . C	59.6	48	$T * A + A * C + T * C$	59.6	48
T . C	52.9 ##	24	$T * A + A * C$	112.5 ##	72
T . A	49.5 ##	10	$A * C + T * C$	109.1 ##	58
A . C	14.1	10	$T * C + T * A$	73.7	58

L. R. S. : Likelihood Ratio Statistic  
 D. F. : Degrees of Freedom  
 # :  $p < 0.05$   
 ## :  $p < 0.01$

相互作用 T . C は有意に大きくはなく、両都市の抗体分布に差があるとは言えない。但し、サンプルサイズが小さく、従って検出力も小さいから同じであると結論するには検討が必要である。また、Table 8 は Samarinda, Surabaya, Kolaka, Jayapura の 4 都市の解析

Table 5 Expected values based on model  $T * A + A * C + T * C$  in table 4

City	Age	Antibody					
		$\leq 4 \times$	$8 \times$	$16 \times$	$32 \times$	$64 \times$	$128 \times$
Vientien	0-9	0.37	3.42	8.21	9.71	7.51	4.77
	10-19	0.32	1.85	5.55	12.9	6.97	2.39
	20-	0.31	1.73	19.2	40.4	28.5	10.8
Samarinda	0-9	0	5.08	5.43	3.94	3.37	2.18
	10-19	0	5.37	7.16	10.2	6.11	2.13
	20-	0	4.54	22.4	28.8	2225	8.70
Kolaka	0-9	0.28	4.26	3.46	5.07	4.69	4.24
	10-19	0.33	3.19	3.23	9.31	6.02	2.93
	20-	0.39	3.55	13.3	34.6	29.3	15.8
Surabaya	0-9	0.43	3.72	6.24	3.88	3.14	2.59
	10-19	0.69	3.78	7.92	9.70	5.47	2.43
	20-	0.87	4.49	34.8	38.4	28.4	14.0
Lombok	0-9	0.28	4.24	5.80	4.23	2.73	2.73
	10-19	0.37	3.59	6.13	8.81	3.97	2.14
	20-	0.35	3.17	20.1	26.0	15.3	9.14
Jayapura	0-9	1.63	3.28	6.85	3.17	1.57	2.50
	10-19	3.29	4.22	11.01	10.0	3.47	2.98
	20-	2.07	2.50	24.1	19.8	8.96	8.52

結果であるが、ここでは抗体、都市の 2 因子交互作用は有意でないものの、その L. R. S. は比較的大きい。これは Jayapura の抗体分布が他の 3 都市と少し異なっているためと思われ、この点を検討する目的で Table 9 の解析を行った。この解析では全体を 6 元の分割表として取り扱っており、6 個の因子 T, A, C1, C2, C3, C4 の内 C1, C2, C3, C4 が各々

1000 のセルに Samarinda のデータを、

0100 のセルに Surabaya のデータを、

0010 のセルに Lombok のデータを、

0001 のセルに Jayapura のデータを、

いれてある。それ以外のセルは全て観察しえないセル、即ち、実験的ゼロのセルとして取り扱った。詳しくは Fienberg<sup>2)</sup>を参照されたい。Table 9 の 2-5 のモデルは各々の都市を次のように取り扱っていることになる。

2 Samarinda vs others

3 Surabaya vs others

4 Lombok vs others

Table 6 Calculation of odds ratio for the  $2 \times 2$  table made from Table 5 (age : 0-9, 10-19, antibody :  $8 \times$ ,  $32 \times$ )

Vientien	$3.42 \times 1.85 \times 9.71 \times 12.9 = 2.46$
Samarinda	$5.08 \times 5.37 \times 3.94 \times 10.2 = 2.46$
Kolaka	$4.26 \times 3.19 \times 5.07 \times 9.31 = 2.45$
Surabaya	$3.72 \times 3.78 \times 3.88 \times 9.70 = 2.46$
Lombok	$4.24 \times 3.59 \times 4.23 \times 8.81 = 2.46$
Jayapura	$3.28 \times 4.22 \times 3.17 \times 10.0 = 2.46$

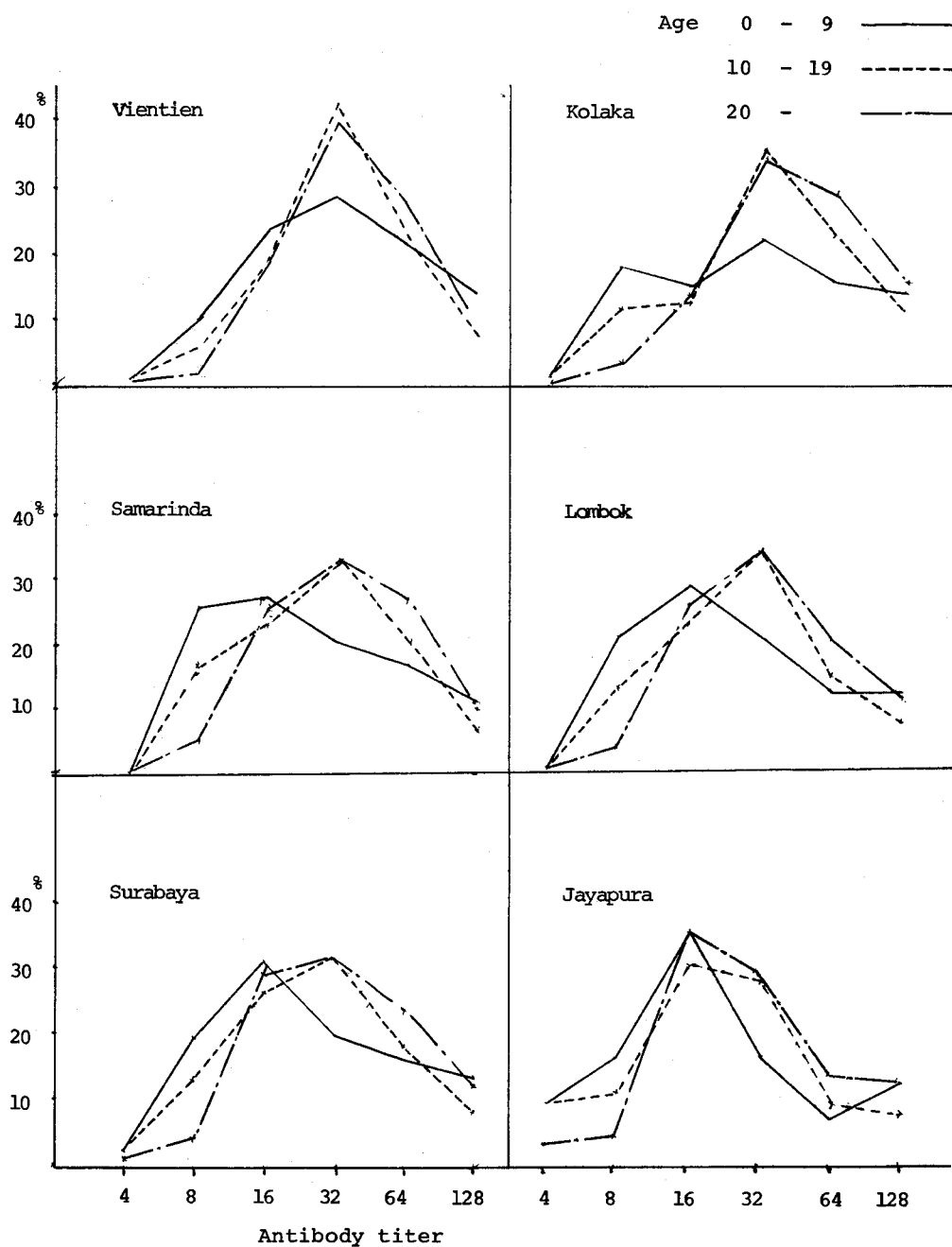


Fig. 1 Age distribution of antibody based on expected values in table 5

### 5 Jayapura vs others

解析結果をみると上の4つの中では5のモデルの適合度が最も良く、2因子交互作用T.C4もL.R.S.=16.6, 自由度=5で有意に大きい。従って、Jayapuraを除く他の3都市の抗体分布は似ているが、Jayapura

のみは異なっていると結論される。

次に東南アジア6都市と札幌市の成績をあわせて解析したところ、3因子交互作用T.A.CはL.R.S.=88.6, 自由度=58と有意に大きく、飽和モデル以外は当てはめられなかった。

Table 7 Loglinear analysis of the data in Vientien and Kolaka

Effect	L. R. S.	D. F.	Model	L. R. S.	D. F.
T . A . C	7.12	10	$T * A + A * C + T * C$	7.12	10
T . C	6.34	5	$T * A + A * C$	13.46	15
T . A	16.3	10	$A * C + T * C$	23.4	20
L. R. S. : Likelihood Ratio Statistics			D. F. : Degrees of Freedom		

Table 8 Loglinear analysis of the data in Samarinda, Surabaya, Lombok and Jayapura

Effect	L. R. S.	D. F.	Model	L. R. S.	D. F.
T . A . C	40.2	28	$T * A + A * C + T * C$	40.2	28
T . C	21.4	14	$T * A + A * C$	61.7	42
T . A	45.4##	10	$A * C + T * C$	85.6##	38
L. R. S. : Likelihood Ratio Statistics			D. F. : Degrees of Freedom		
# : $p < 0.05$			## : $p < 0.01$		

Table 9 Loglinear analysis of the data in Jayapura, Lombok, Samarinda, and Surabaya

Factor		
T : Titer	$\leq 4 \times / 8 \times / 32 \times / 64 \times / 128 \times$	
A : Age	0-9 / 10-19 / 20-	
C1	1/0 : Samarinda vs others	
C2	1/0 : Surabaya vs others	
C3	1/0 : Lombok vs others	
C4	1/0 ; Jayapura vs others	
Model	L. R. S.	D. F.
1 $T * A + A * C1 * C2 * C3 * C4 + T * C1 * C2 * C3 * C4$	40.2	28
2 " + $T * C1$	52.2	38
3 " + $T * C2$	60.5#	37
4 " + $T * C3$	59.9#	37
5 " + $T * C4$	45.0	37
6 $T * A + A * C1 * C2 * C3 * C4$	61.7	42

# :  $p < 0.05$ 

#### 4 考 察

東南アジア 6 都市の年齢別抗ロタウイルス中和抗体価を表わす 3 元分割表は 3 因子交互作用のないモデルで説明することが可能であった。6 都市の抗体分布に差はみられるものの、年齢の上昇に伴う抗体分布の変化には共通の傾向があり、Table 6 のような関係と考えられる。しかし、これが具体的にどのような意味を持つかを説明するのは難しい。一般的に言って、反応（回帰分析に於ける目的変数、ここでは抗体価）が 3 つ以上のカテゴリーに分れるときは、2 つの場合と違って 3

因子交互作用の意味を解釈するのは困難である。札幌市の抗体分布を含めたとき 3 因子交互作用が有意に大きかったことについても同じ理由で解釈は難しいが、この場合のように都市間の抗体分布に大きな差のあるとき、対数線型解析を行うのは不適当であろうと思われる。

東南アジア 6 都市の抗ロタウイルス中和抗体価の分布は対数線型解析の結果から、Vientien, Kolaka からなる群, Samarinda, Surabaya, Lombok からなる群および Jayapura の 3 つに分けられ、またロタウイルスの侵いん度もこの順に低くなっている。各都市の地理

的分布は Fig. 2 に示す通りであって、Samarinda, Surabaya, Lombok の 3 都市は互いに近くに位置している。しかし Kolaka はこれら 3 都市に近いにも関わらず、その抗体分布はむしろ遠く離れた Vientien のそれに類似している。Kolaka の位置するセレベス島が開発の遅れた地域であることと関連しているかもしれない。

上に述べたような各都市の抗体分布の違いは生のデータを詳細に検討することでも発見可能であるが、本報のごとく適当なモデルをあてはめて得られた期待値を検討することにより、もっと容易にみつかることができる。また主観的に分類した結果を Table 7-9 のように客観的に評価することも可能である。

対数線型解析は、ここに示したような例だけでなく、分割表の解析に一般に広く応用でき、今後更に普及するものと思われる。ただ、この解析にはコンピュータが必要であり、これが多くの研究者にとって障害のひとつとなっているであろうことは容易に想像される。我々の開発した対数線型解析プログラムはこの障害を取り除く上で有用であると思われる。使用方法も簡単であり、試みに本学の公衆衛生、衛生学実習で学生に使用させたところ、短時間の内に使用方法を習得させることができた。なお、パーソナルコンピュータを用い

ているため、計算結果の精度、計算に要する時間などが問題になるが、本報で扱ったセルの数が 100 程度の分割表の解析であれば実用上、問題ないと思われる。

## 5 結 論

パーソナルコンピュータによる対数線型解析プログラムを開発し、これを用いて、東南アジア 6 都市及び日本（札幌）における年齢別抗ロタウイルス中和抗体価の解析を行い、以下の結果を得た。

- 1 抗体価、年齢、都市の 3 因子を用いて解析したところ、東南アジア 6 都市の年齢別抗体分布に 3 因子交互作用のないモデルを適合させることが分った。
- 2 このモデルを利用して、東南アジア 6 都市の抗体分布を 3 つに分類できた。
- 3 札幌市における抗体分布は東南アジア 6 都市のそれと大きく異なっており、対数線型モデルによる解析を行うことはできなかった。
- 4 パーソナルコンピュータによる対数線型解析プログラムは使用法も簡単で疫学研究などに有用と思われる。

## 文 献

1. Urasawa, S., Urasawa, T., Djoko, Y., Furuya, K., Akiba, S., Kanamitsu, M. : A survey of rotavirus infection in the tropics. *Jpn. J. Med. Sci. Biol.* **34**, 293-298 (1981).
2. Fienberg, S. E. : The analysis of cross-classified categorical data. MIT Press, London (1980).
3. Upton, G. : The analysis of cross-tabulated data. John Wiley & Sons, New York (1978).
4. Anderson, S., Auquier, A., Hauck, W. W., Oakes, D., Vandaele, W., Weisberg, H. I. : Statistical methods for comparative studies. John Wiley & Sons, New York (1981).
5. 石居進 : 生物統計学入門. 培風館, 東京 (1975).
6. Baker, R., Nelder, J. : The GLIM system release 3 manual. The Numerical Algorithms Group, Oxford (1978).

別刷請求先 :

(〒060) 札幌市中央区南 1 条西 17 丁目

札幌医科大学衛生学講座 秋葉澄伯

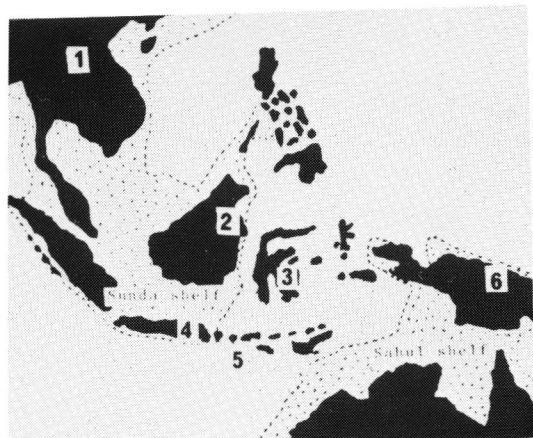


Fig. 2 Location of six cities

- |             |            |
|-------------|------------|
| 1 VIENTIANE | 4 SURABAYA |
| 2 SAMARINDA | 5 LOMBOK   |
| 3 KOLAKA    | 6 JAYAPURA |